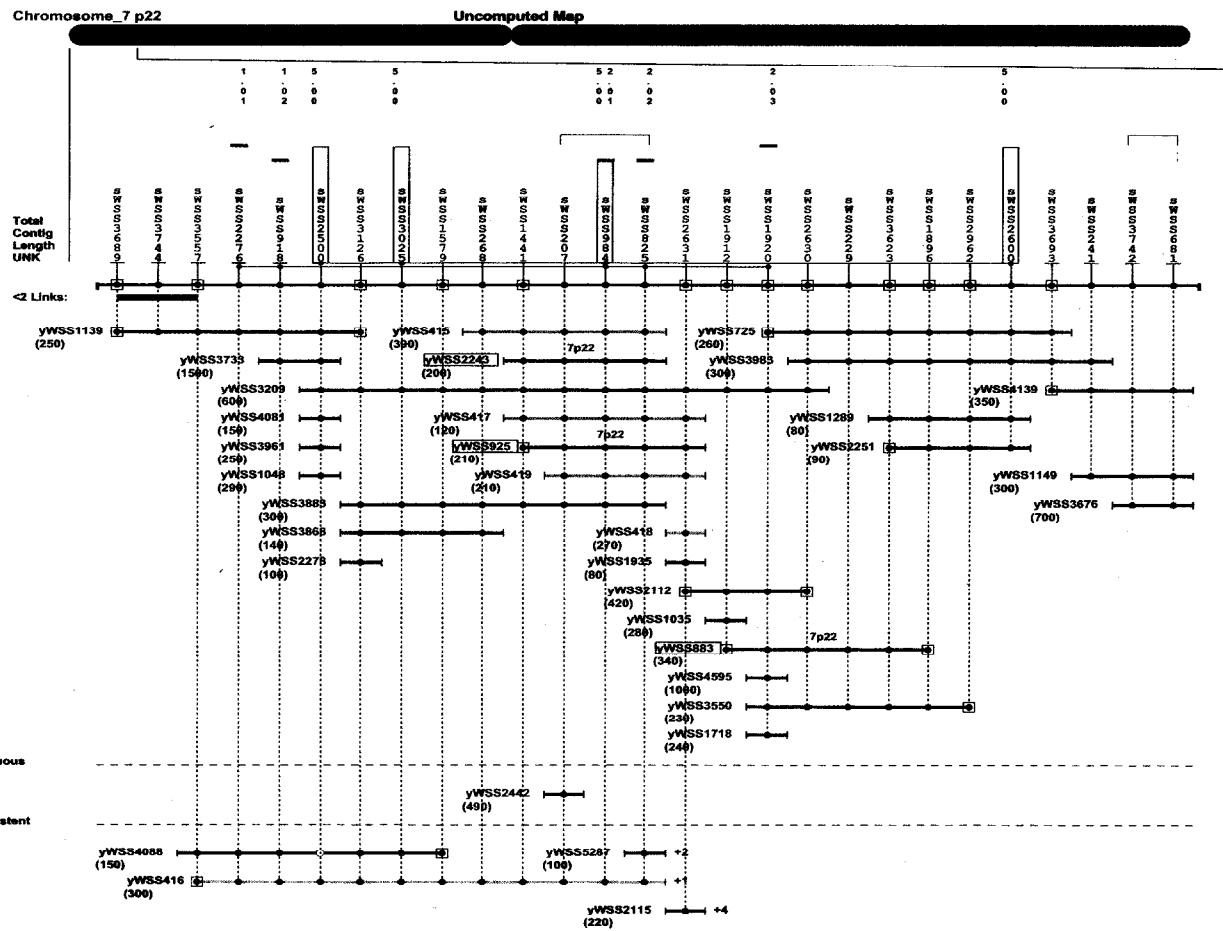


# CLONE-BASED PHYSICAL MAPPING



Eric D. Green, M.D., Ph.D.  
Chief, Genome Technology Branch  
National Human Genome Research Institute

Phone: 301-402-0201  
FAX: 301-402-4735  
E-Mail: [egreen@nhgri.nih.gov](mailto:egreen@nhgri.nih.gov)

# **Clone-Based Physical Mapping**

## **I. Overview of Physical Maps**

## **II. Large DNA Cloning Systems**

- A. Cosmids (Fosmids)**
- B. P1s**
- C. PACs**
- D. BACs**
- E. YACs**

## **III. Strategies for Physical Mapping**

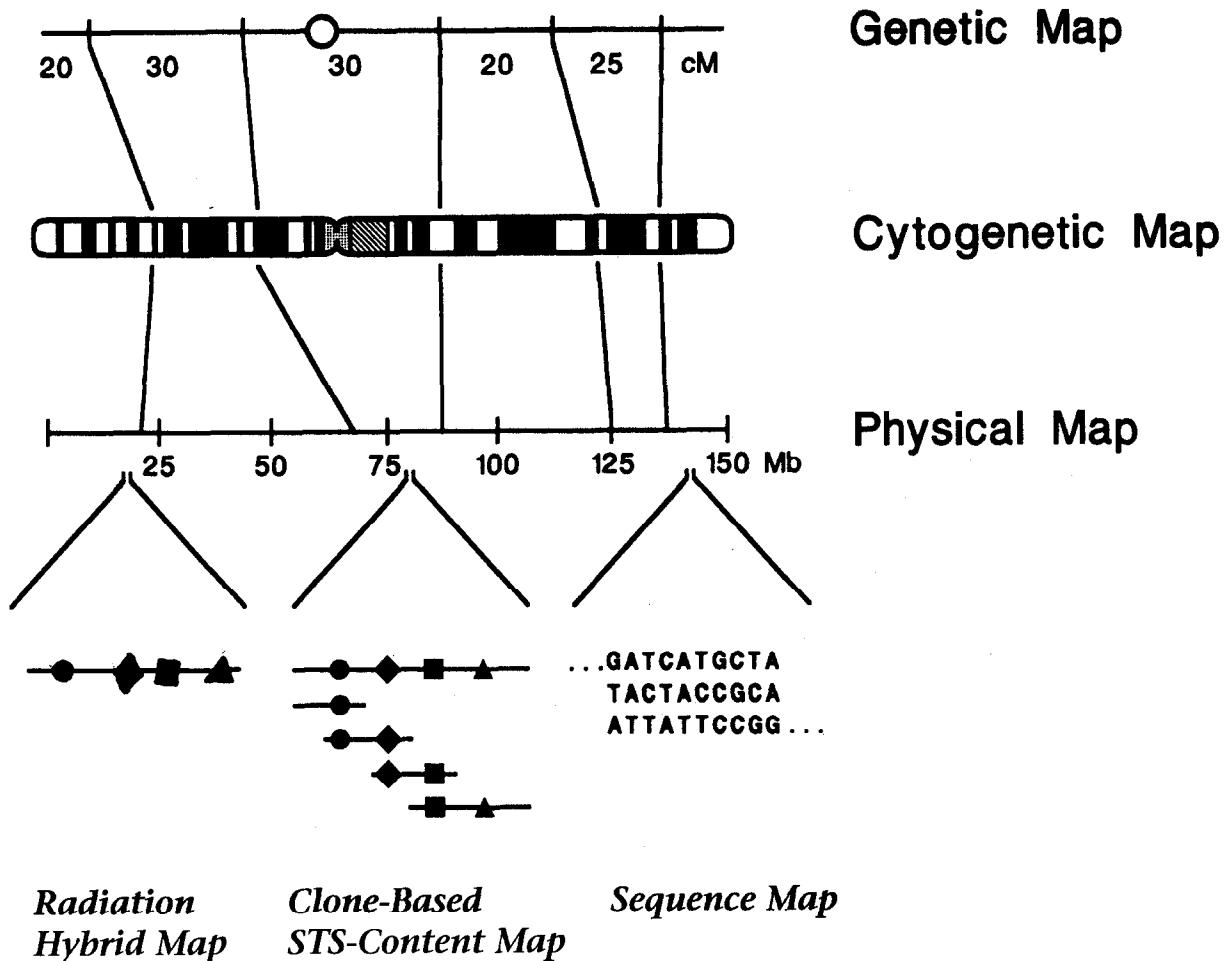
## **IV. Status of Clone-Based Physical Maps of Mammalian Genome**

## **V. Future Prospects**

### **General Plan for Lecture:**

- Stress the Practical Aspects of Physical Mapping
- Focus on the Mapping of Mammalian Genomes (Mostly Human)
- Highlight Relevant Literature
- Provide Information on Relevant Electronic Resources

# Types of Genomic Maps



Importance of Physical Maps:

- Localization and Isolation of Genes (e.g., Positional Cloning)
- Study of Genome Organization and Evolution
- Framework for Systematic DNA Sequencing

# Role of Physical Mapping in Positional Cloning

[Collins (1992), Ballabio (1993), Collins (1995)]

## Positional Cloning of Human Disease Genes

### Pedigree Collection

- Phenotype Determination
- DNA Collection



### Genetic Mapping

- Localize to Chromosomal Region
- Refined Localization



### Physical Mapping and Clone Isolation



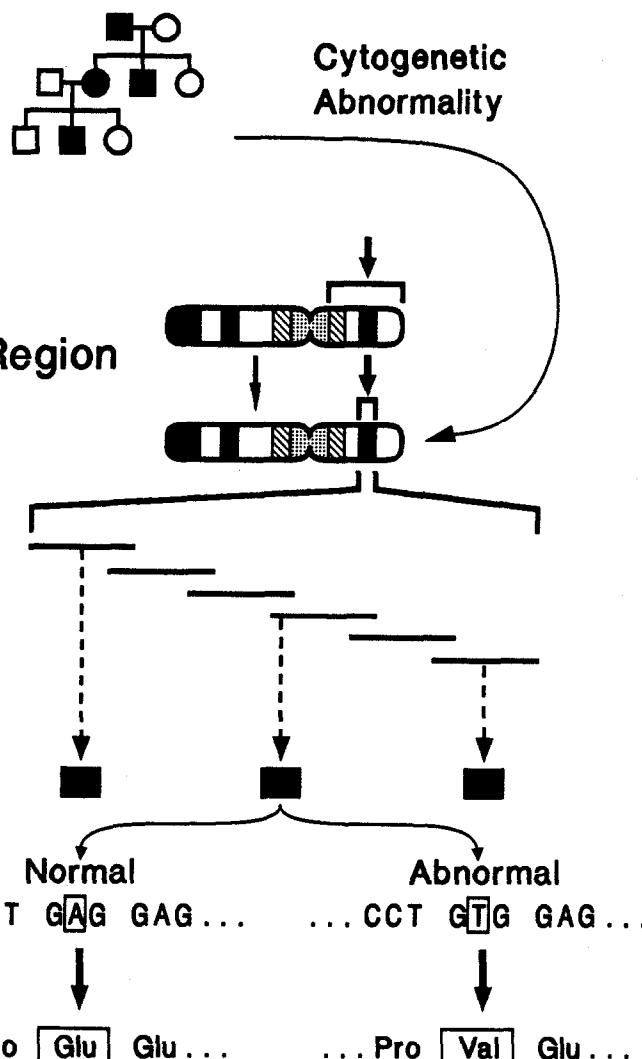
### Gene Isolation



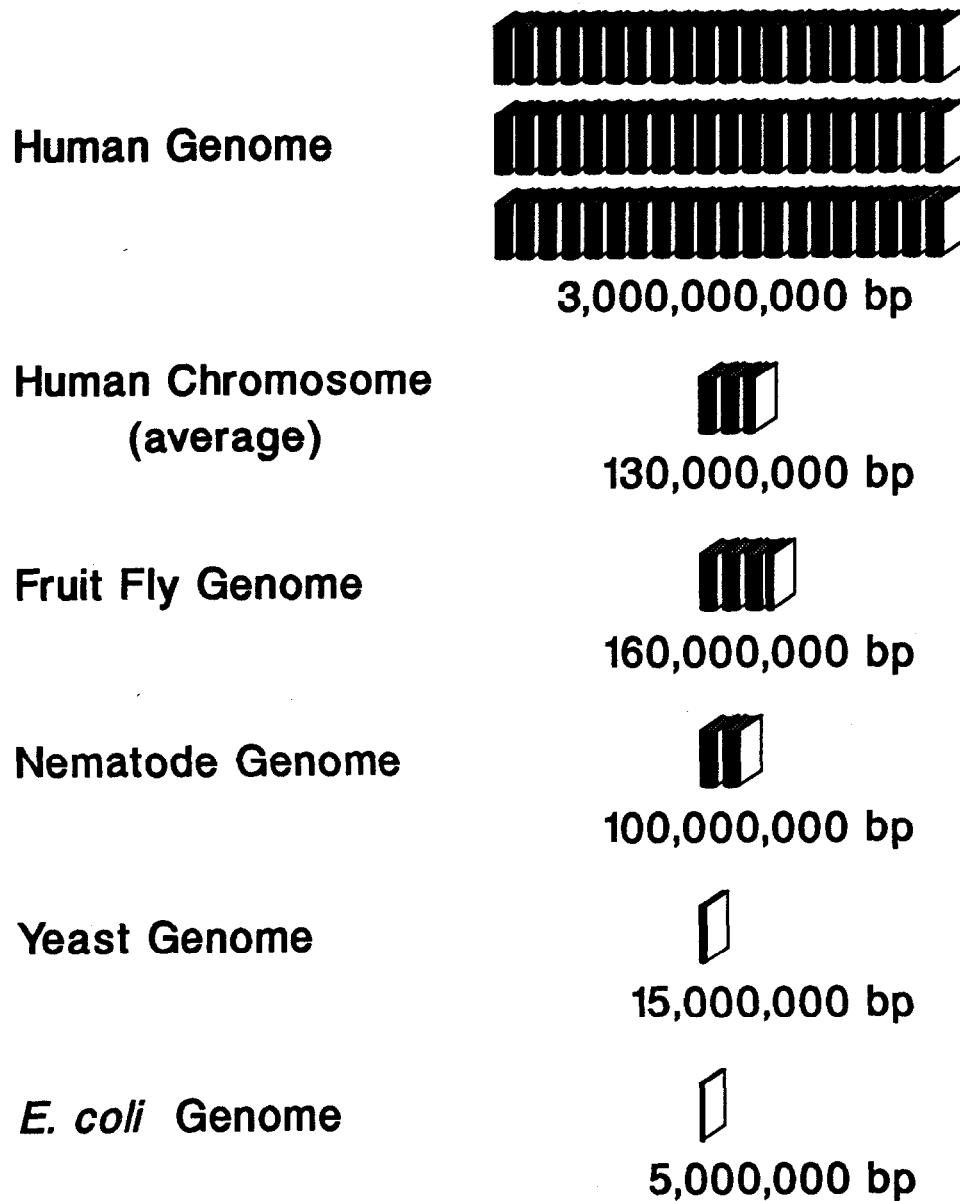
### Mutation Analysis



### Functional Studies



# Genome Sizes



# Fundamentals of Physical Mapping

- “Mapping is About Order”

-*Maynard Olson* (1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997...)

- Physical Mapping Involves:

Ordering of Clones and/or Landmarks

Typically with Some Physically Measurable Metric

- Examples of ‘Landmark Only’ Ordering:

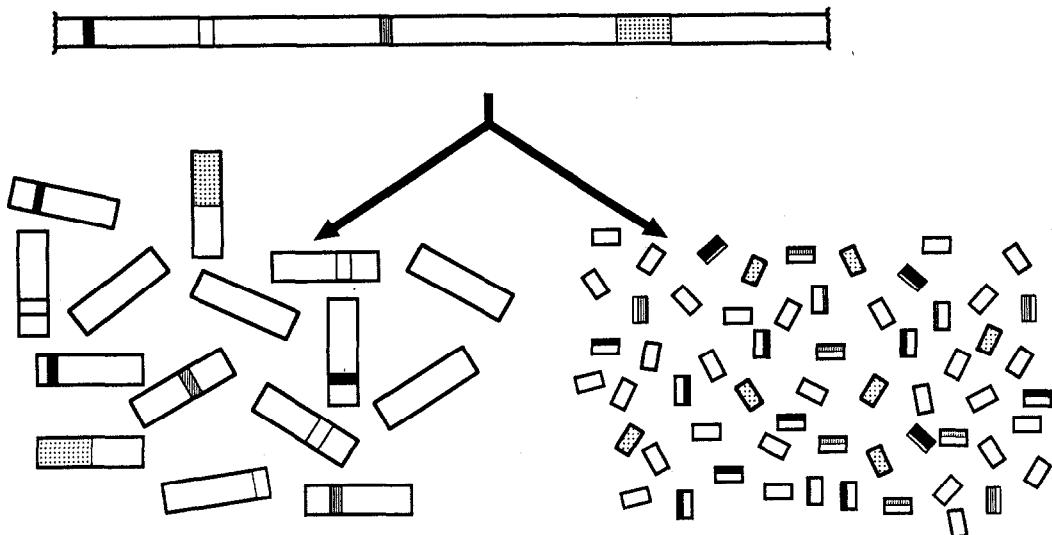
Long-Range Restriction Mapping by Pulsed-Field Gel Electrophoresis

Radiation Hybrid (RH) Mapping

- Theoretical Discussions of Clone-Based Physical Mapping

Arratia et al. (1991), Barillot et al. (1991),  
Palazzolo et al. (1991), Olson and Green (1993)

- Clone-Based Physical Mapping: ‘Jigsaw Puzzle Analogy’



# **Large DNA Cloning Systems**

- Reasonably Arbitrary Definition of 'Large'
- Bacterial- vs. Yeast-Based Host Systems
- Want the Cloned DNA to Accurately Reflect the Starting Genome

Problem of Instability

Problem of Chimerism

- Development of 'Array Mentality' for Clone Libraries

Clones Arrayed in Individual Wells of Microtiter Plates

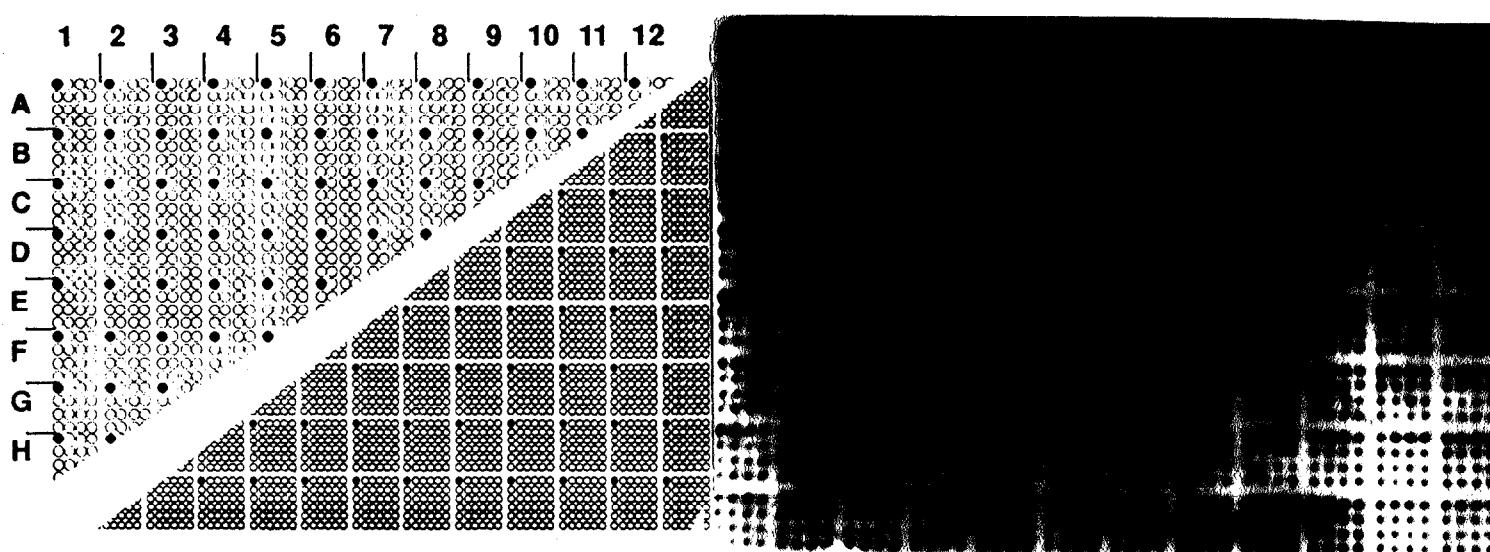
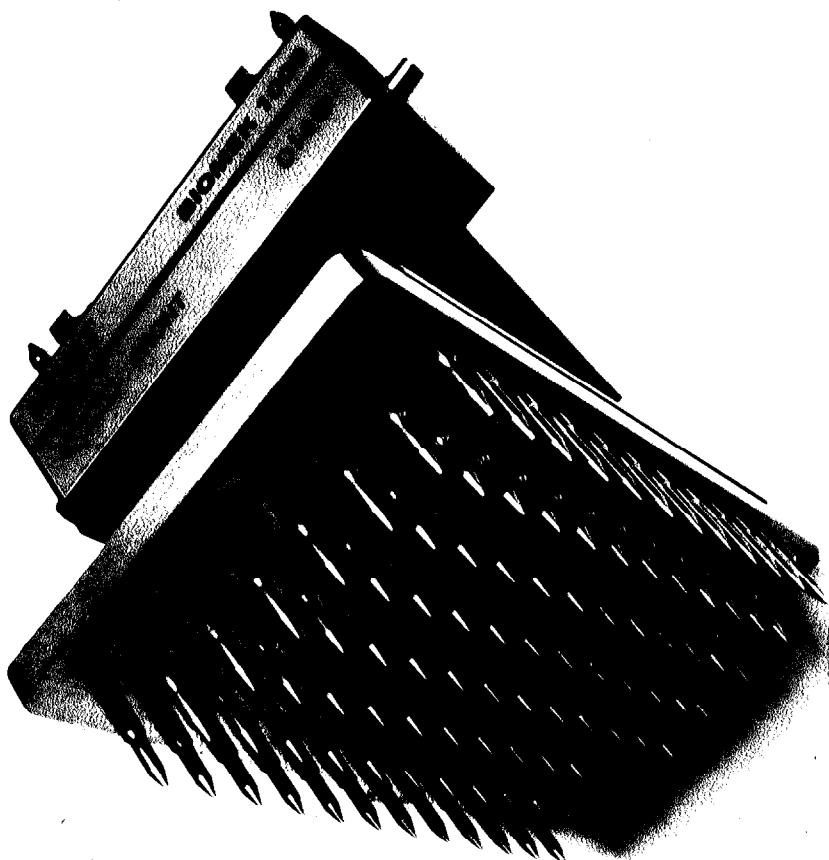
Various Densities Available (96-, 192-, and 384-Well Plates)

- Advantages of Arrayed Clone Resources ('Reference Libraries')

1. Simplicity of Storing and Transferring Clone Collections
2. Repeated Hybridization-Based Screening
3. Repeated PCR-Based Screening
4. Convenient Format for Retrieving Clones of Interest
5. Ability to Assimilate Data on Common Clones

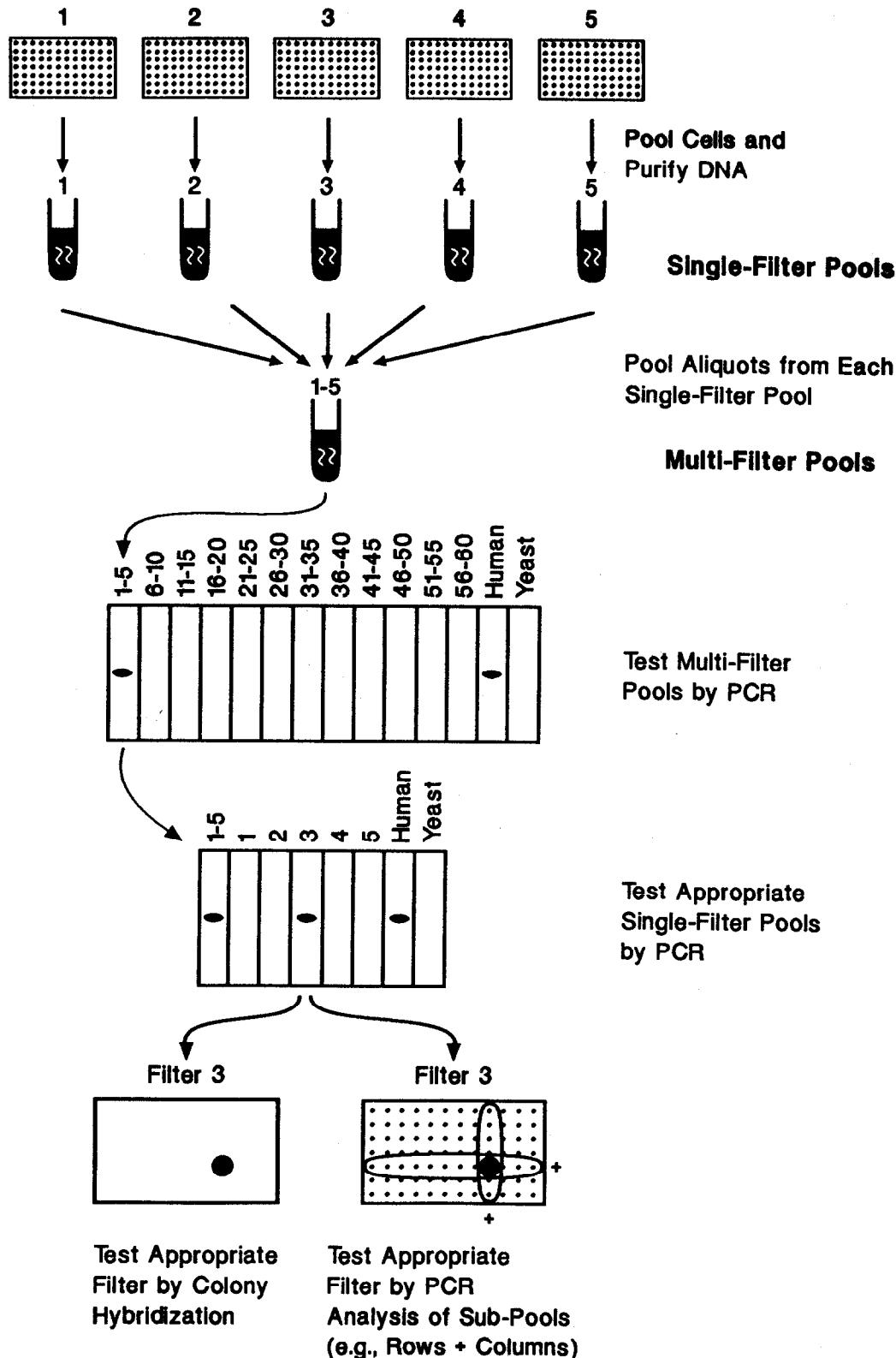
# Hybridization-Based Analysis of High-Density Arrays

Bentley et al. (1992), Ross et al. (1992), Olsen et al. (1993)



# PCR-Based Analysis of Arrayed Libraries

Green and Olson (1990)



## **Commercial Involvement in Clone Distribution**

Research Genetics, Inc.

800-533-4363  
205-533-4363  
<http://www.resgen.com>

Genome Systems, Inc.

800-430-0030  
314-692-0033  
<http://www.genomesystems.com>

ATCC

800-638-6597  
301-881-2600  
<http://www.atcc.org>

# Cosmids

- Bacterial-Based Cloning System
- ‘Antique’ of the Large DNA Cloning Systems
- Plasmid Vector with Bacteriophage Packaging Sequences (*cos* Sites)
- High Efficiency Packaging System

Relatively Homogeneous Insert Sizes

Libraries from Small Amounts of DNA (e.g., Flow-Sorted DNA)

Antibiotic Selection

- Cloned Inserts: 35-45 kb, Circular DNA
- High Copy Number

High Yields of DNA by Standard Methods

Instability Problems (Despite Recombination-Deficient Hosts)

- Relatively Non-Chimeric
- Various Libraries (Whole Genomes, Individual Chromosomes)
- References

Sambrook et al. (1989), Wahl et al. (1987), Nizetic et al. (1991),  
Evans et al. (1992), Ivens et al. (1993)

- ‘Fosmids’ [Kim et al. (1992)]

Cosmid Vector Engineered with F Factor

Low Copy → More Stable

# P1 Clones

- Bacterial-Based Cloning System
- Developed by Sternberg (1990)

## **Bacteriophage P1 cloning system for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs**

(DNA packaging/pac cleavage/genome mapping/gene isolation)

NAT STERNBERG

- Bacteriophage P1 (Genome Size: 100 kb)
- P1-Based Vector and Complex P1 Packaging Extracts

Limited to 100 kb (Constraints of Viral Particle)  
2 loxP Sites Results in Circularization of DNA  
Antibiotic Selection

- Cloned Inserts: 70-100 kb, Circular DNA
- Low Copy Number

Low Yields of DNA by Standard Methods  
Highly Stable (with Recombination-Deficient Hosts)  
Potential for IPTG Induction → 10-30 Fold Increase

- Relatively Non-Chimeric
- Human and Mouse Libraries Commercially Available
- References

Sternberg (1990), Sternberg et al. (1990),  
Pierce and Sternberg (1992), Pierce and Sternberg (1992),  
Shepherd et al. (1994)

# P1-Derived Artificial Chromosomes (PACs)

- Bacterial-Based Cloning System
- Developed by Ioannou et al. (1994)

## A new bacteriophage P1-derived vector for the propagation of large human DNA fragments

Panayiotis A. Ioannou<sup>1</sup>, Chris T. Amemiya<sup>1</sup>, Jeffrey Garnes<sup>1</sup>, Peter M. Kroisel<sup>1</sup>, Hiroaki Shizuya<sup>2</sup>, Chira Chen<sup>1,3</sup>, Mark A. Batzer<sup>1</sup> & Pieter J. de Jong<sup>1,3</sup>

- Slightly Modified P1 Vector

Lacks Packaging Signal  
Antibiotic Selection

- Transform by Electroporation
- No Packaging of DNA → Larger Size Capacity
- Cloned Inserts: 100-150 kb, Circular DNA
- Low Copy Number

Low Yields of DNA by Standard Methods  
Highly Stable (with Recombination-Deficient Hosts)

- Relatively Non-Chimeric
- Human and Mouse Libraries Commercially Available

# Bacterial Artificial Chromosomes (BACs)

- Bacterial-Based Cloning System
- Developed by Shizuya et al. (1992)

## **Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector**

(electroporation/physical mapping/human genome)

**HIROAKI SHIZUYA\*, BRUCE BIRREN, UNG-JIN KIM, VALERIA MANCINO, TATIANA SLEPAK,  
YOSHIKI TACHIKI, AND MELVIN SIMON†**

- Based on the *E. coli* F Factor (Fertility Plasmid): Replication Control
- BAC Vectors

Cloning site in LacZ Gene (Blue/White Selection)  
Antibiotic Selection

- Transform by Electroporation
- No Packaging of DNA → Larger Size Capacity
- Cloned Inserts: 100-150 kb, Circular DNA
- Low Copy Number

Low Yields of DNA by Standard Methods  
Highly Stable (with Recombination-Deficient Hosts)

- Relatively Non-Chimeric
- Human and Mouse Libraries Commercially Available

# **Yeast Artificial Chromosomes (YACs)**

- Developed by Burke et al. (1987)

## **Cloning of Large Segments of Exogenous DNA into Yeast by Means of Artificial Chromosome Vectors**

**DAVID T. BURKE, GEORGES F. CARLE, MAYNARD V. OLSON**

- Yeast-Based Cloning System

*Saccharomyces cerevisiae* as Host (Eukaryotic Cell)  
More Hospitable Environment for Eukaryotic DNA (?)

- System Based on Ability to 'Harness' Cloned DNA with the Structural Elements Required for Propagation of a Linear Chromosome in Yeast
- Cloned Insert: ~100 to >1,000 kb, Linear DNA
- Spheroplast Transformation Procedure

Technically Demanding  
Poorly Defined Upper Size Limit for Cloned Insert

- References

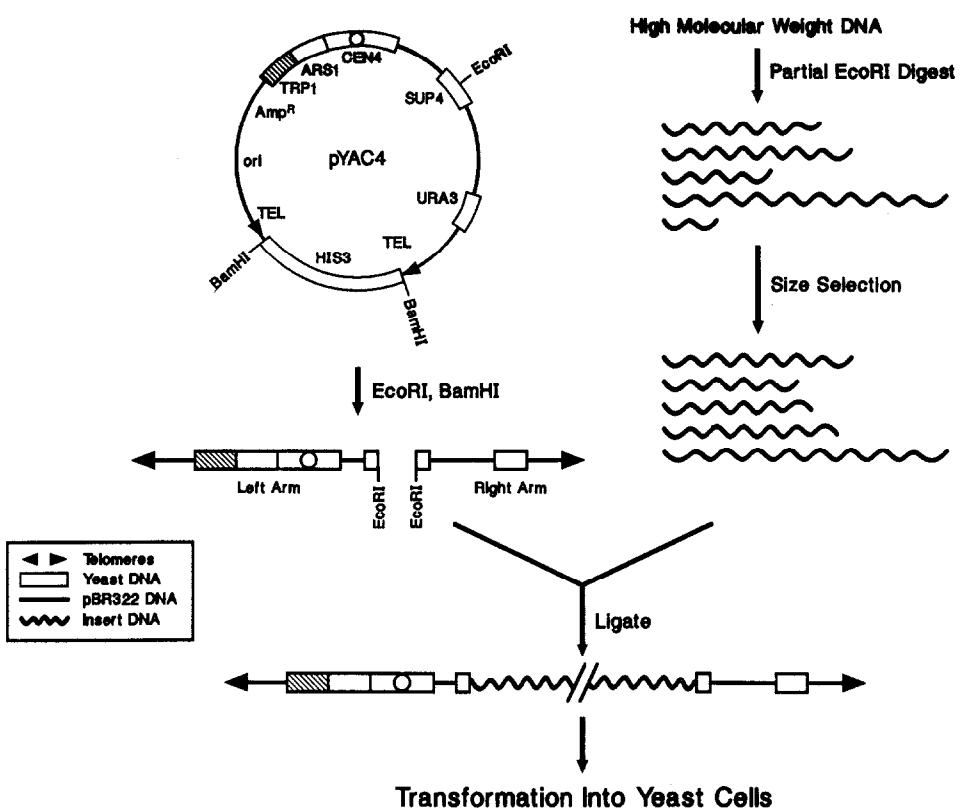
Hieter et al. (1990), Schlessinger (1990), Burke (1991),  
Ramsay and Wicking (1991), Schlessinger and Kere (1992)

# YAC Cloning System

- YAC Vector

- 2 Telomere Sequences
- 1 Centromere Sequence
- 1 ARS Sequence
- 2 Yeast Selectable Markers (Not Antibiotic Selection)

- Strategy for YAC Cloning



# **Major Features of YACs**

- Cloned DNA in Single Copy within Yeast Genome

Generally Same Structure and Size as Endogenous Chromosomes  
Limited ‘Access’ to Cloned DNA (e.g., Gel Isolation)

- Chimerism as Major ‘Problem’

Upwards of 40-60% of Clones in Total Mammalian DNA Libraries  
Largely from ‘Overly’ Efficient Yeast Recombination System  
[Green et al. (1991)]

- Instability (e.g., Internal Deletions) as More Minor ‘Problem’

Difficult to Estimate Extent of Problem  
Clearly Related to Recombination in Yeast  
Might be Related to YAC Size  
Influenced by Clone Handling

- Issues of Recombination-Deficient Host Strains

- Various Human, Mouse, and Rat Libraries Constructed

## **Human:**

**Washington University** [Burke and Olson (1991), Brownstein et al. (1989)]  
**CEPH** (Includes ‘Mega-YACs’) [Albertsen et al. (1990), Dausset et al. (1992)]  
**ICRF** [Larin et al. (1991)]  
**ICI** [Anand et al. (1989), Anand et al. (1990)]

## **Mouse:**

**Princeton** [Burke et al. (1991), Rossi et al. (1992)]  
**St. Mary’s** [Chartier et al. (1992)]  
**ICRF** [Larin et al. (1991, 1993)]  
**Whitehead** [Kusumi et al. (1993), Haldi et al. (1996)]

## **Rat:**

**Harvard** [Cai et al. (1997)]  
**Whitehead** [Haldi et al. (1997, 1997)]

# Strategies for Clone-Based Physical Mapping

- Two Key Components ('Jigsaw Puzzle Analogy')

Cloned Fragments (Pieces of the Puzzle)

Landmarks (Provide Clues for Fitting Pieces Together)

- Ultimately Want to Order Clones and/or Landmarks
- Ideally Want 'Access' to Both Clones and Landmarks
- Clone-Based Physical Mapping Typically Involves the Use of Landmarks to Assembly Clone 'Contigs'

**Contig:** Overlapping Set of Clones that Together Contain a Contiguous Segment of the Source Genome

- Nature of Landmarks

Must Provide 'Unique' Information About the DNA

Must be Easy to Identify

Can be Intrinsic or Extrinsic to the Clones

- Early Candidates for Landmarks: Restriction Sites (Intrinsic Landmarks)

## **Physical Maps of “Smaller” Genomes**

- *E. Coli* [Kohara et al. (1987)]

### **The Physical Map of the Whole *E. coli* Chromosome: Application of a New Strategy for Rapid Analysis and Sorting of a Large Genomic Library**

Yoji Kohara,\* Kiyotaka Aida,\*, and Kazumi Isono†

- Yeast [Olson et al. (1986), Riles et al. (1993)]

### **Random-clone strategy for genomic restriction mapping in yeast**

(DNA/chromosomes/λ cloning/algorithm)

MAYNARD V. OLSON, JAMES E. DUTCHIK, MADGE Y. GRAHAM, GARRETT M. BRODEUR, CYNTHIA HELMS,  
MARK FRANK, MIA MACCOLLIN, ROBERT SCHEINMAN, AND THOMAS FRANK

### **Physical Maps of the Six Smallest Chromosomes of *Saccharomyces cerevisiae* at a Resolution of 2.6 Kilobase Pairs**

Linda Riles,\* James E. Dutchik,<sup>†,1</sup> Amara Baktha,\* Brigid K. McCauley,<sup>\*,2</sup> Edward C. Thayer,<sup>\*,3</sup>  
Mary P. Leckie,<sup>†,4</sup> Valerie V. Braden,\* Julie E. Depke<sup>†,5</sup> and Maynard V. Olson<sup>†,6</sup>

- Nematode [Coulson et al. (1986)]

### **Toward a physical map of the genome of the nematode *Caenorhabditis elegans***

(ordered clone bank/genomic data base/clone matching)

ALAN COULSON, JOHN SULSTON, SYDNEY BRENNER, AND JONATHAN KARN

# Early Physical Mapping of Human Chromosomes

- Strategies Analogous to that Used with *E. coli*, Yeast, and Nematode Applied to Several Human Chromosomes

Cosmid Clones (e.g., Flow-Sorted Libraries)

Restriction Map Construction and/or Fingerprint Analysis  
[e.g., Stallings et al. (1990)]

Difficult to Construct Long-Range Contiguous Maps

- Shift in Strategies with the Development of YACs
- Modified Fingerprint-Based Strategies Attempted with YACs
- Distinguishing Features of YACs

No Ability to Readily Purify Cloned DNA

Fingerprint Analysis:

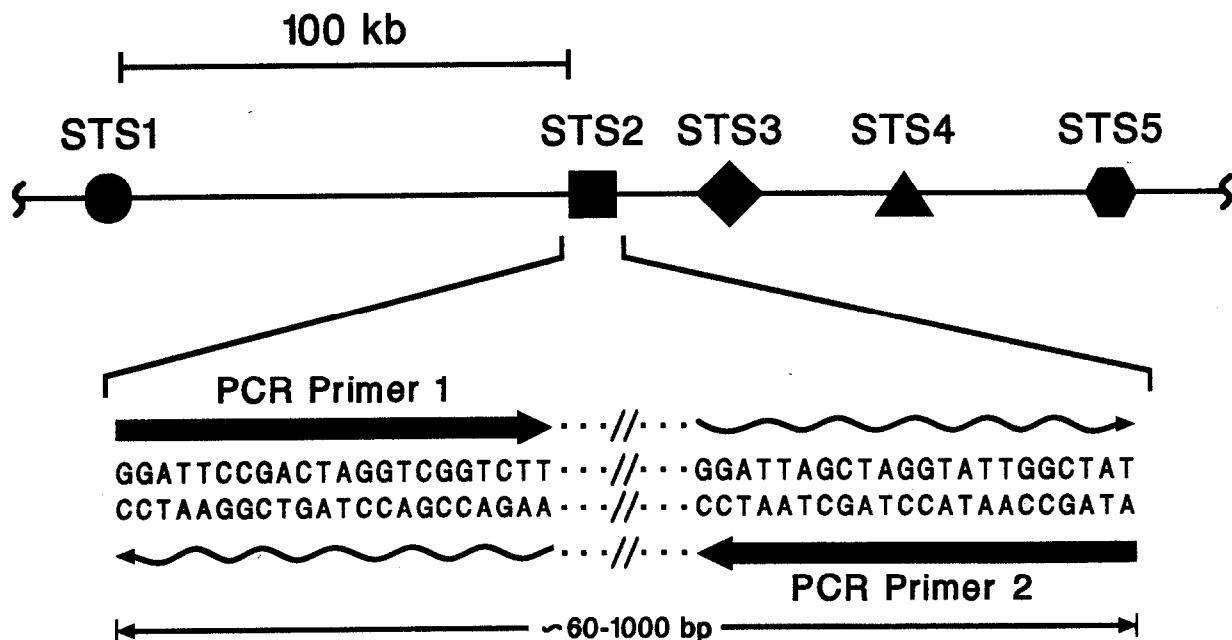
1. Typically Requires Gel-Transfer Hybridization
2. Typically Uses Repetitive Element-Specific Probe(s)

Establish YAC ‘Fingerprint’ → Infer Overlap(s) with Other YACs

- Fingerprint-Based Approaches Provide Clone-Based Maps Without Deriving Extrinsic Landmarks
- Major ‘Evolution’ in Strategy Occurred with Development of PCR

## Sequence-Tagged Sites (STSs)

- Development of PCR → Profound Impact on Physical Mapping
- STSs as Landmarks for Constructing Physical Maps

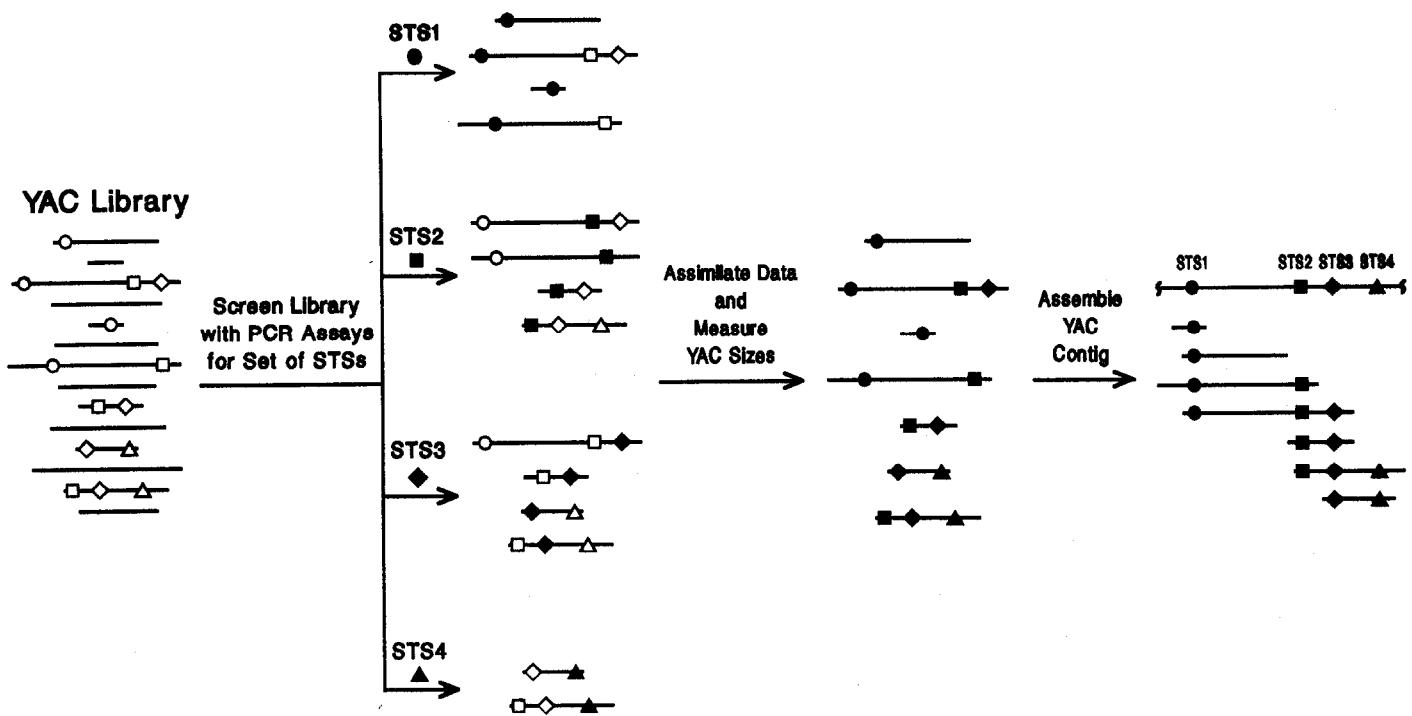


- “Common Language” Proposal by Olson et al. (1989)

### A Common Language for Physical Mapping of the Human Genome

MAYNARD OLSON, LEROY HOOD,  
CHARLES CANTOR, DAVID BOTSTEIN

# YAC-Based STS-Content Mapping



- Green and Olson (1990) as Successful Demonstration of Paradigm

## Chromosomal Region of the Cystic Fibrosis Gene in Yeast Artificial Chromosomes: A Model for Human Genome Mapping

ERIC D. GREEN AND MAYNARD V. OLSON

# **Implications of STSs as Landmarks for Physical Maps**

- STSs are Extrinsic to the Clones
- Advantages of STSs as Landmarks

PCR-Based (Sensitivity, Specificity, Automation)  
Electronic-Based ‘Transfer’ of STSs  
Landmarks Are Independent of the Mapping Resource  
Sequence-Based Nature Facilitates Integration with Sequence Map

- General Review on STS-Content Mapping: Green and Green (1991)
- Programmatic Goal of U.S. Human Genome Project

100-kb Average Resolution STS Map of Human Genome  
[Collins and Galas (1993)]  
Therefore, ~30,000 STSs for Human Genome

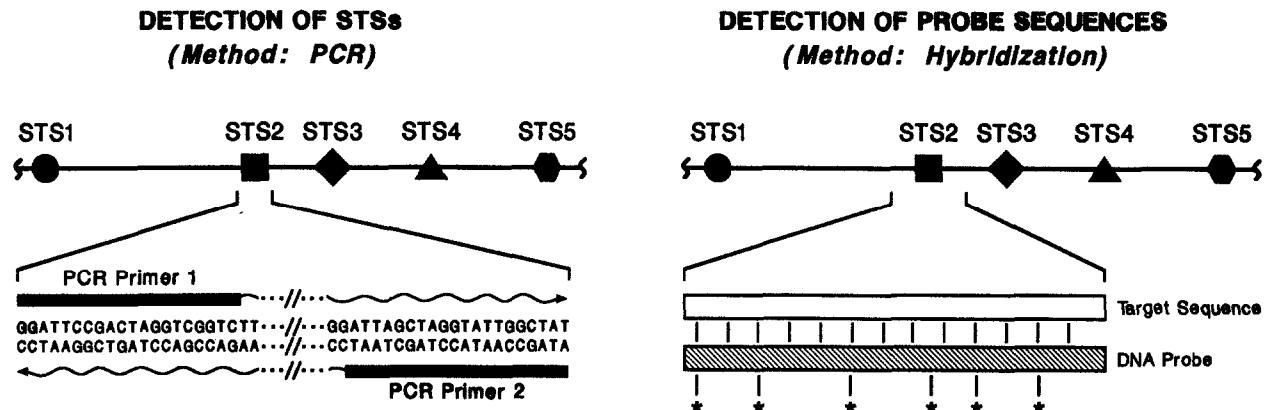
- YAC-Based STS Map as ‘Intermediate Map’ En Route to Sequencing

## **Development of STSs**

- Operational Definition of an STS
  1. Sequence That Can be Amplified by a PCR Assay
  2. Functionally is ‘Unique’ in the Genome
- DNA Sequence → Select Primers → Confirm Above Definition
- Generation of Sequences for Developing STSs
  1. Non-Targeted (i.e., Genome Wide)
  2. Targeted
- Targeted Approaches
  - Specific Chromosomes
    - Somatic Hybrid Cell Lines (Subcloning, PCR)
    - Flow Sorting
    - Microdissection
  - Genetic Markers (Microsatellites)
  - Expressed Sequences (Genes, ESTs)
    - Note: Sequences from 3' Ends of Genes Preferred for STS Development (Since Less Likely to Contain an Intron or to be Present in Other Related Genes)

# Conceptual Similarity of STSs and Probes

- Detection of STS- vs. Probe-Based Landmarks



- Continued Interest in Probe-Based Map Construction

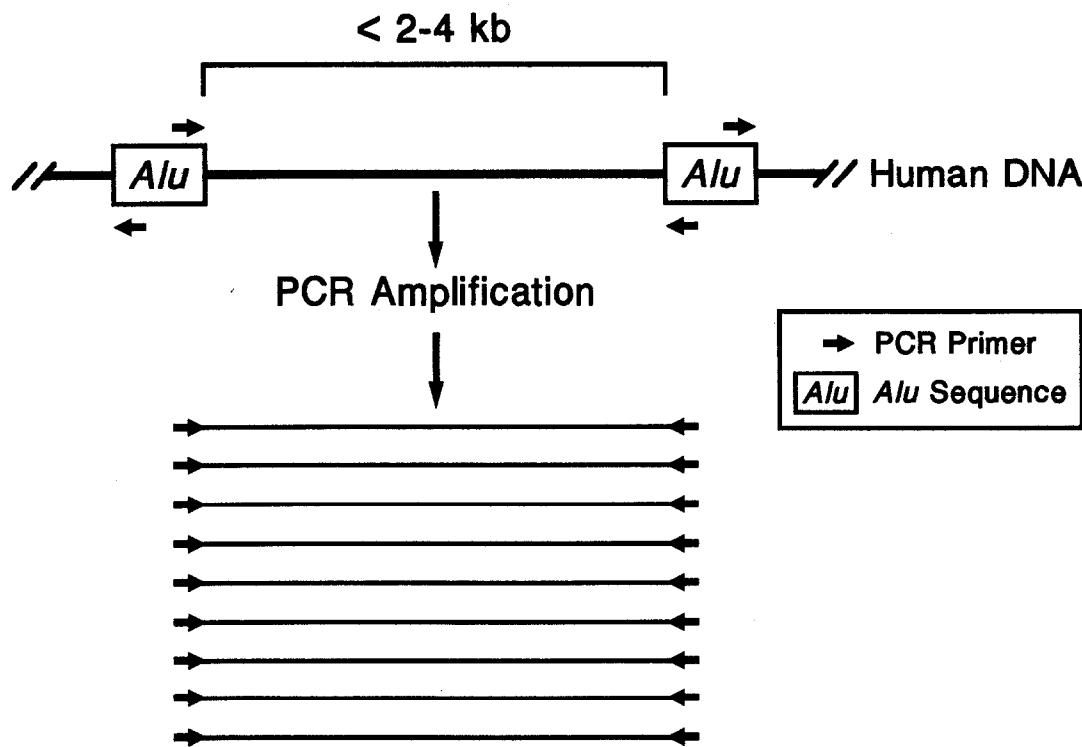
Hybridization of Restriction Digested YAC DNA → Fingerprints

Hybridization of Arrayed YAC Colonies or DNA

# Interspersed Repetitive Sequence (IRS) PCR

- Near Ubiquitous Presence of Repetitive Sequences in Mammalian DNA
- Amplification from Priming Sites within Closely-Spaced Repeat Units
- Primers Designed to “Highly Conserved” Regions of Repeat Units
- Most Extensively Utilized IRS-PCR Method: “Alu PCR”

Nelson et al. (1989), Nelson et al. (1991), Nelson (1991)



# **Clone-Based Physical Mapping of the Human Genome**

- YACs for ‘1<sup>st</sup> Generation’ Clone-Based Physical Maps
- Options for Landmarks:
  - Fingerprints
  - Hybridization Probes (e.g., Alu-PCR Products)
  - STSs
- Strategic Issue of Genome-Wide vs. Chromosome-Specific Efforts
- 1<sup>st</sup> Generation Clone-Based Physical Maps of Human Chromosomes:
  - Strategies for Construction
  - Characteristics
  - Lessons Learned
  - Overall Progress to Date
  - Future Prospects

# **Early Successes in Human Chromosome Mapping**

- Chromosome Y [Foote et al. (1992), Vollrath et al. (1992)]

## **The Human Y Chromosome: Overlapping DNA Clones Spanning the Euchromatic Region**

**Simon Foote, Douglas Vollrath, Adrienne Hilton, David C. Page**

## **The Human Y Chromosome: A 43-Interval Map Based on Naturally Occurring Deletions**

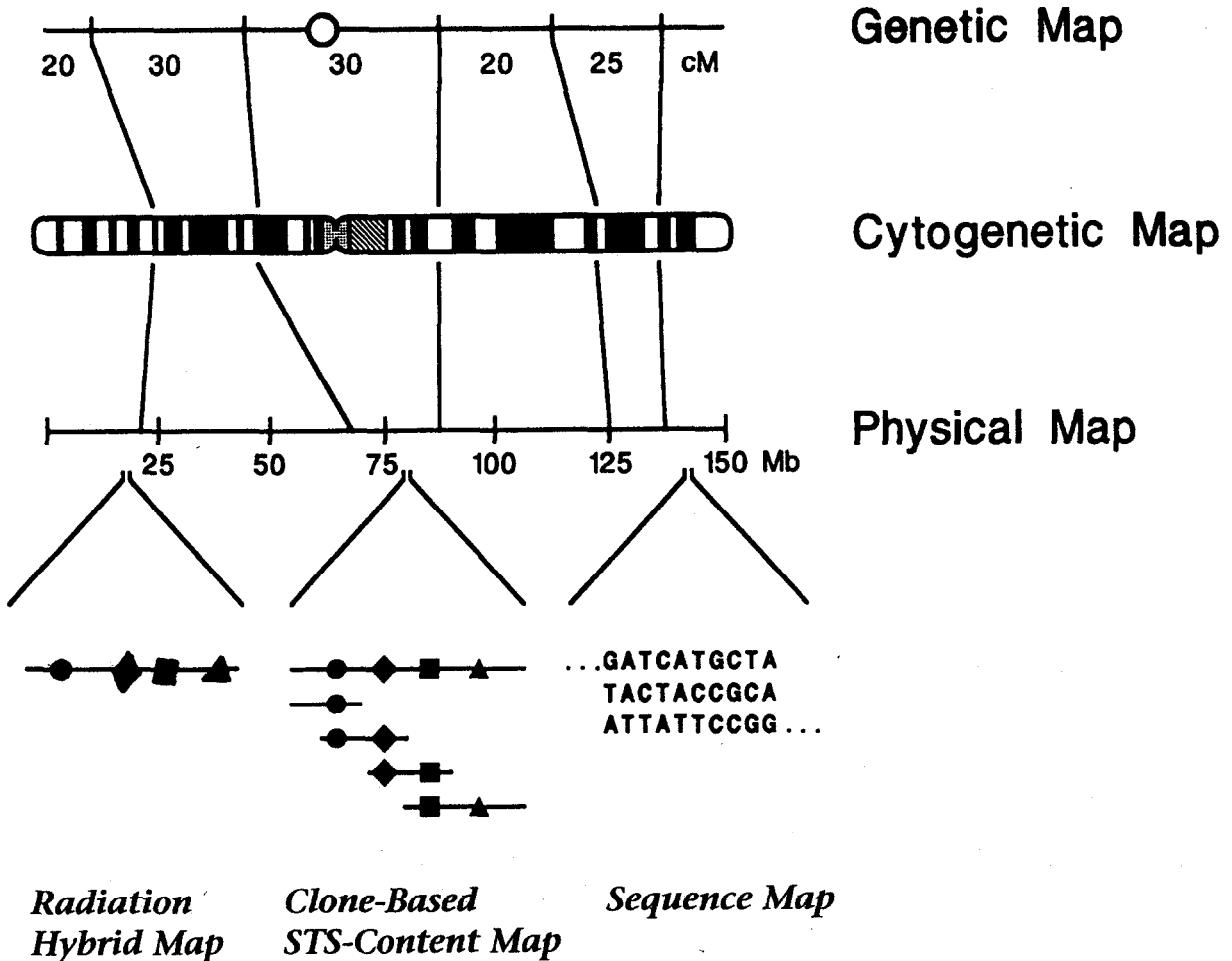
**Douglas Vollrath, Simon Foote, Adrienne Hilton, Laura G. Brown,  
Peggy Beer-Romero, Jonathan S. Bogan, David C. Page**

- Chromosome 21 [Chumakov et al. (1992)]

## **Continuum of overlapping clones spanning the entire human chromosome 21q**

**Ilya Chumakov<sup>\*†</sup>, Philippe Rigault<sup>†</sup>, Sophie Guillou<sup>†</sup>, Pierre Ougen<sup>\*</sup>, Alain Billaut<sup>\*</sup>,  
Ghislaine Guasconi<sup>†</sup>, Patricia Gervy<sup>†</sup>, Isabelle LeGall<sup>\*</sup>, Pascal Soubarue<sup>\*</sup>, Laurent Grinias<sup>†</sup>,  
Lydie Bougueret<sup>\*</sup>, Christine Bellanné-Chantelot<sup>\*</sup>, Bruno Lacroix<sup>\*</sup>, Emmanuel Barillot<sup>†</sup>,  
Philippe Gesnouin<sup>†</sup>, Stuart Pook<sup>†</sup>, Guy Vaysseix<sup>\*†</sup>, Gerard Frelat<sup>‡</sup>, Annette Schmitz<sup>‡</sup>,  
Jean-Luc Sambucy<sup>\*</sup>, Assumpcio Bosch<sup>§</sup>, Xavier Estivill<sup>§</sup>, Jean Weissenbach<sup>†||</sup>, Alain Vignal<sup>†</sup>,  
Harold Riethman<sup>¶</sup>, David Cox<sup>\*</sup>, David Patterson<sup>\*\*</sup>, Kathleen Gardiner<sup>\*\*</sup>, Masahira Hattori<sup>††</sup>,  
Yoshiyuki Sakaki<sup>††</sup>, Hitoshi Ichikawa<sup>††</sup>, Misao Ohki<sup>††</sup>, Denis Le Paslier<sup>\*</sup>, Roland Hellrig<sup>††</sup>,  
Stylianos Antonarakis<sup>††</sup> & Daniel Cohen<sup>\*††</sup>**

# Integration of Genomic Maps



- Importance of Map Integration

Enhances Utility of Each Map

Facilitates the Construction of Maps (Especially Physical Map)

- Strategies for Map Integration

Physical → Cytogenetic

Genetic → Physical

(∴ Genetic → Cytogenetic)

- Importance for Both Genome-Wide and Chromosome-Specific Efforts

# The CEPH/Genethon Map of the Human Genome

- Bellanne-Chantelot et al. (1992)

## Mapping the Whole Human Genome by Fingerprinting Yeast Artificial Chromosomes

Christine Bellanné-Chantelot,<sup>\*</sup> Bruno Lacroix,<sup>\*</sup>  
Pierre Ougen,<sup>\*</sup> Alain Billault,<sup>\*</sup> Sandrine Beaufils,<sup>†</sup>  
Stéphane Bertrand,<sup>†</sup> Isabelle Georges,<sup>†</sup>  
Fabrice Gilbert,<sup>†</sup> Isabelle Gros,<sup>†</sup>  
Georges Lucotte,<sup>†</sup> Laurent Susini,<sup>†</sup>  
Jean-Jacques Codani,<sup>‡</sup> Philippe Gesnouin,<sup>†</sup>  
Stuart Pook,<sup>†</sup> Guy Vaysseix,<sup>\*</sup> Jennifer Lu-Kuo,<sup>§</sup>  
Thomas Ried,<sup>§</sup> David Ward,<sup>§</sup> Ilya Chumakov,<sup>\*</sup>  
Denis Le Paslier,<sup>\*</sup> Emmanuel Barillot,<sup>†</sup>  
and Daniel Cohen<sup>\*†</sup>

- Cohen et al. (1993)

## A first-generation physical map of the human genome

D. Cohen<sup>\*†</sup>, I. Chumakov<sup>\*†</sup> & J. Weissenbach<sup>†‡</sup>

- Chumakov et al. (1995)

## A YAC contig map of the human genome

Ilya M. Chumakov<sup>\*</sup>, Philippe Rigault<sup>\*</sup>, Isabelle Le Gall<sup>\*</sup>, Christine Bellanné-Chantelot<sup>\*</sup>, Alain Billault<sup>\*</sup>, Sophie Guillou<sup>\*</sup>, Pascal Soularue<sup>\*</sup>, Ghislaine Quasconi<sup>\*</sup>, Eric Pouiller<sup>\*</sup>, Isabelle Gros<sup>\*</sup>, Maria Belova<sup>\*</sup>, Jean-Luc Sambucy<sup>\*</sup>, Laurent Susini<sup>\*</sup>, Patricia Gervy<sup>\*</sup>, Fabrice Gilbert<sup>\*</sup>, Sandrine Beaufils<sup>\*</sup>, Hung Bui<sup>\*</sup>, Catherine Massart<sup>\*</sup>, Marie-France De Tand<sup>\*</sup>, Frédérique Dukasz<sup>\*</sup>, Sandrine Lecoulant<sup>\*</sup>, Pierre Ougen<sup>\*</sup>, Virginie Perrot<sup>\*</sup>, Martial Saumier<sup>\*</sup>, Catherine Soravito<sup>\*</sup>, Rita Bahouayila<sup>\*</sup>, Annick Cohen-Akenine<sup>\*</sup>, Emmanuel Barillot<sup>\*</sup>, Stéphane Bertrand<sup>†</sup>, Jean-Jacques Codani<sup>†</sup>, Dominique Caterina<sup>\*</sup>, Isabelle Georges<sup>†</sup>, Bruno Lacroix<sup>\*</sup>, Georges Lucotte<sup>\*</sup>, Mourad Sahbatou<sup>\*</sup>, Christian Schmit<sup>\*</sup>, Muriel Sangouard<sup>\*</sup>, Emmanuel Tubacher<sup>\*</sup>, Colette Dib<sup>†</sup>, Sabine Fauré<sup>†</sup>, Cécile Flizames<sup>†</sup>, Gabor Gyapay<sup>†</sup>, Philippe Millasseau<sup>†</sup>, Simon NGuyen<sup>†</sup>, Delphine Muselet<sup>†</sup>, Alain Vignal<sup>†</sup>, Jean Morissette<sup>†</sup>, Joan Menninger<sup>†</sup>, Jonathan Lleman<sup>†</sup>, Trushna Desai<sup>†</sup>, Amy Banks<sup>†</sup>, Patricia Bray-Ward<sup>†</sup>, David Ward<sup>†</sup>, Thomas Hudson<sup>§</sup>, Sebastian Gerety<sup>§</sup>, Simon Foote<sup>§</sup>, Lincoln Stein<sup>§</sup>, David C. Page<sup>§||</sup>, Eric S. Lander<sup>§||</sup>, Jean Weissenbach<sup>†</sup>, Denis Le Paslier<sup>\*</sup> & Daniel Cohen<sup>\*</sup>

# CEPH/Genethon Strategy for Map Construction

- Experimental Data Set

Hybridization-Based Fingerprint Analysis (~33,000 YACs, 10-Hit)

Hybridization Analysis (YAC x YAC):

Alu-PCR YAC Probes → Alu-PCR Products from YACs

Alu-PCR Hybridization Assignment of YACs to Chromosomes

FISH-Based Assignment of YACs to Chromosomes

\*\*\*Assignment of Genethon Genetic Markers (STSs) to YACs

- Data Analysis

Complicated!!!

Suite of Programs to ‘Disambiguate’ the Data (*Quickmap*)

Statistical Analysis of Fingerprint Data to Assess ‘Overlaps’

Consideration of YAC x YAC Hybridization Results

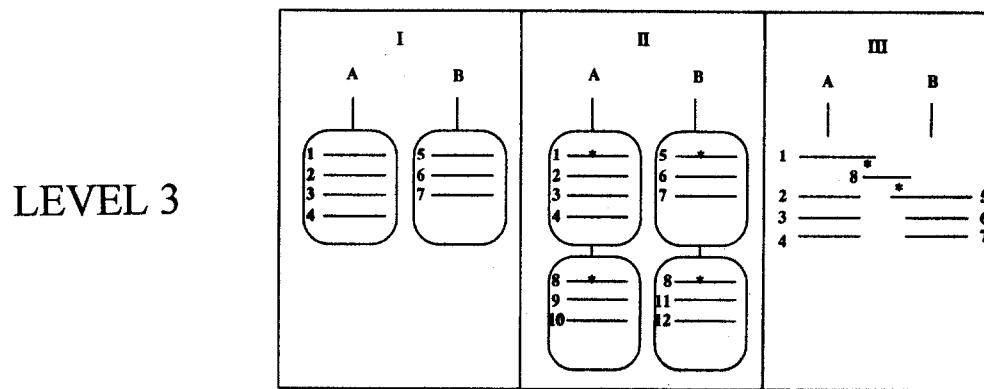
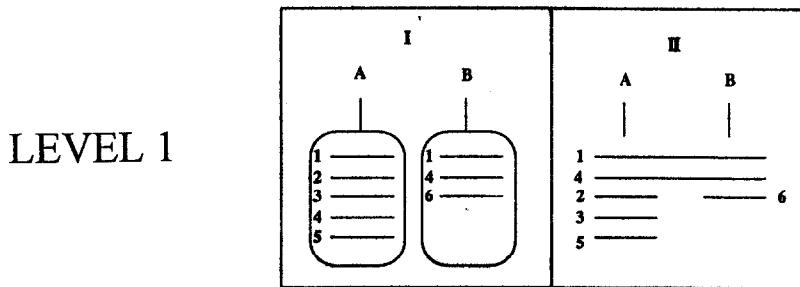
Factor the Chromosomal Assignments of YACs

Heavy Reliance on Genethon Genetic Map for Contig Assembly

Predict ‘Most Likely’ Paths Among Overlapping Clones

# CEPH/Genethon Physical Map

- Path Levels



- Map Highlights

Clone-Based Map (Mostly Based on Intrinsic Landmarks)

225 Contigs Averaging 10 Mb, 75% of Genome Covered

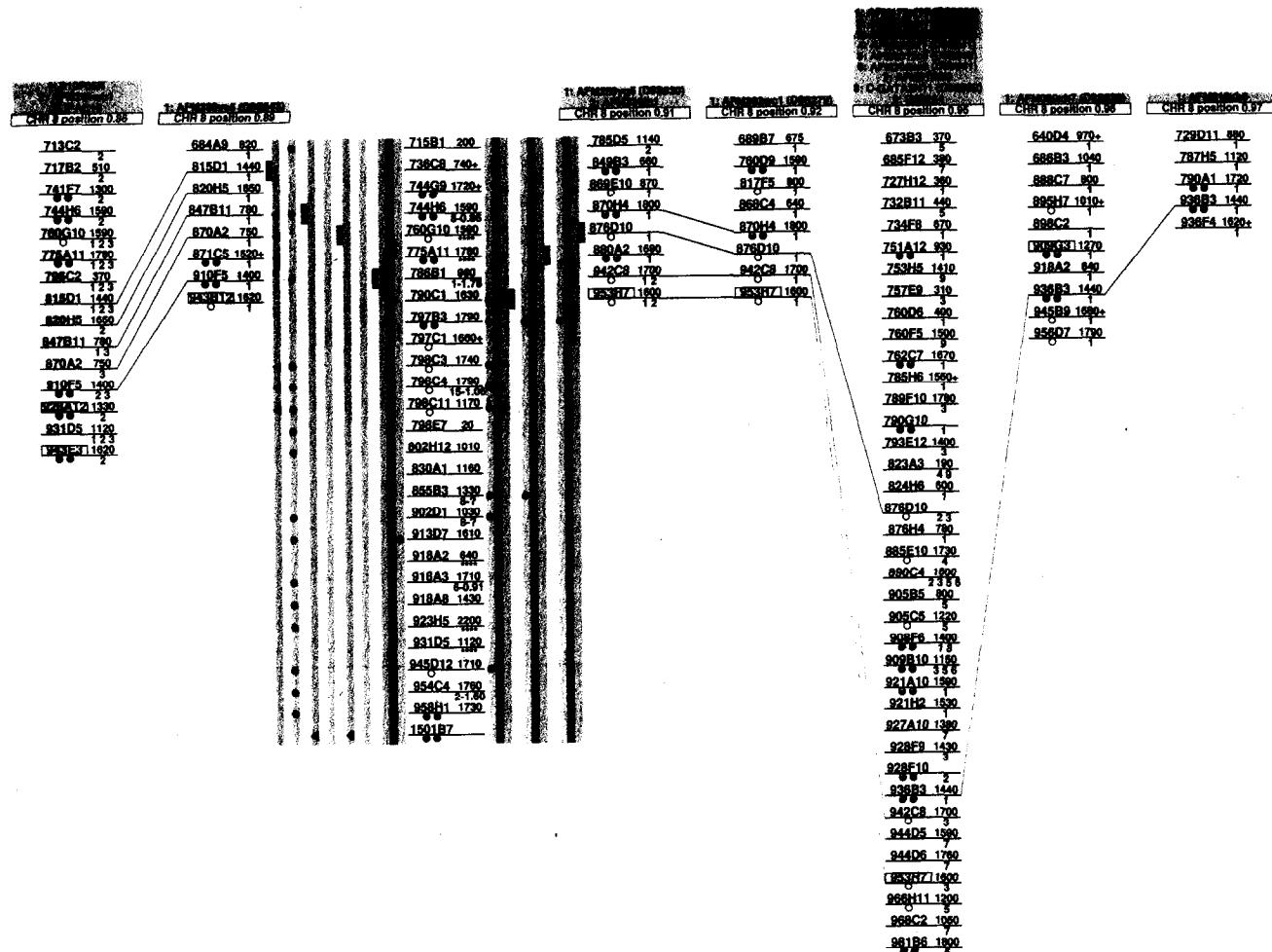
Potentially Useful for Positional Cloning Projects

Poor Scaffold for DNA Sequencing (Sparse STS Density)

- Data and Map Availability

<http://www.cephb.fr/bio/ceph-genethon-map.html>

# Example of a CEPH/Genethon Map



## **Chromosome-Specific Mapping Efforts**

- Chromosome 3 (U. of Colorado, U. of Texas-San Antonio)  
Gemmill et al. (1995)  
<http://www-eri.uchsc.edu/Welcome.html>  
<http://mars.uthscsa.edu/DB>
- Chromosome 4 (Stanford U.)  
Goold et al. (1993)  
<http://www-shgc.stanford.edu>
- Chromosome 7 (NHGRI, U. of Toronto)  
Green et al. (1994, 1995), Bouffard et al. (1997)  
<http://www.nhgri.nih.gov/DIR/GTB/CHR7>  
<http://www.genet.sickkids.on.ca/chromosome7/>
- Chromosome 10 (Genome Therapeutics)  
<http://www.cric.com/htdocs/sequences/chr10-mapping/index.html>
- Chromosome 11 (U. of Texas-Southwestern, Roswell Park)  
Smith et al. (1993), Quackenbush et al. (1995)  
Qin et al. (1996)  
<http://mcdermott.swmed.edu/datapage>  
<http://shows.med.buffalo.edu/home.html>
- Chromosome 12 (Albert Einstein U.)  
Krauter et al. (1995)  
<http://paella.med.yale.edu/chr12/Home.html>

## **Chromosome-Specific Mapping Efforts [con't]**

- Chromosome 16 (Los Alamos Labs)  
Doggett et al. (1995)  
<http://www-ls.lanl.gov>
- Chromosome 19 (Lawrence Livermore Labs)  
Ashworth et al. (1995)  
<http://www-bio.llnl.gov/bbrp/genome/genome.html>
- Chromosome 21 (U. of Colorado, CEPH)  
<http://www-eri.uchsc.edu/Welcome.html>
- Chromosome 22 (U. of Pennsylvania, Sanger Centre)  
Bell et al. (1995)  
Collins et al. (1995)  
<http://www.cbil.upenn.edu/HGC22.html>  
<http://www.sanger.ac.uk>
- Chromosome X (Washington University, Baylor University)  
Nagaraja et al. (1997)  
<http://genome.wustl.edu/cgm/cgm.html>  
<http://gc.bcm.tmc.edu:8088/home.html>
- Chromosome Y  
Foote et al. (1992), Vollrath et al. (1992)  
<http://www-genome.wi.mit.edu>

## **'100-kb STS Maps' of Human Chromosomes**

- Nagaraja et al. (1997)

RESEARCH

## **X Chromosome Map at 75-kb STS Resolution, Revealing Extremes of Recombination and GC Content**

Ramaiah Nagaraja, Sandra MacMillan, Juha Kere, Carmela Jones,  
Stephanie Griffin, Matthew Schmatz, Jennifer Terrell,  
Michael Shomaker, Christopher Jermak, Christian Hott,  
Mochubeloa Masisi, Steven Mumm, Anand Srivastava, Giuseppe Pilia,  
Terence Featherstone, Richard Mazzarella, Sheila Kesterson,  
Brigid McCauley, Brian Railey, Frank Burrough, Volker Nowotny,  
Michele D'Urso, David States, Bernard Brownstein, and  
David Schlessinger<sup>1</sup>

- Bouffard et al. (1997)

RESEARCH

## **A Physical Map of Human Chromosome 7: An Integrated YAC Contig Map with Average STS Spacing of 79 kb**

Gerard G. Bouffard,<sup>1</sup> Jacquelyn R. Idol,<sup>1</sup> Valerie V. Braden,<sup>1</sup>  
Leslie M. Iyer,<sup>1</sup> Aimee F. Cunningham,<sup>1</sup> Lauren A. Weintraub,<sup>1</sup>  
Jeffrey W. Touchman,<sup>1</sup> Rose M. Mohr-Tidwell,<sup>2</sup> Dale C. Peluso,<sup>2</sup>  
Robert S. Fulton,<sup>2</sup> Melinda S. Ueltzen,<sup>2</sup> Jean Weissenbach,<sup>3</sup>  
Charles L. Magness,<sup>4</sup> and Eric D. Green<sup>1,5</sup>

# The Whitehead Map of the Human Genome

- Hudson et al. (1995)

## An STS-Based Map of the Human Genome

Thomas J. Hudson,\* Lincoln D. Stein, Sebastian S. Gerety, Junli Ma, Andrew B. Castle, James Silva, Donna K. Slonim, Rafael Baptista, Leonid Kruglyak, Shu-Hua Xu, Xintong Hu, Angela M. E. Colbert, Carl Rosenberg, Mary Pat Reeve-Daly, Steve Rozen, Lester Hui, Xiaoyun Wu, Christina Vestergaard, Kimberly M. Wilson, Jane S. Bae, Shanak Maitra, Soula Ganiatsas, Cheryl A. Evans, Margaret M. DeAngelis, Kimberly A. Ingalls, Robert W. Nahf, Lloyd T. Horton Jr., Michele Oskin Anderson, Alville J. Collymore, Wenjuan Ye, Vardoukie Kouyoumjian, Irena S. Zemsteva, James Tam, Richard Devine, Dorothy F. Courtney, Michelle Turner Renaud, Huy Nguyen, Tara J. O'Connor, Cécile Fizames, Sabine Fauré, Gabor Gyapay, Colette Dib, Jean Morissette, James B. Orlin, Bruce W. Birren, Nathan Goodman, Jean Weissenbach, Trevor L. Hawkins, Simon Foote, David C. Page, Eric S. Lander\*

- Web Site: <http://www-genome.wi.mit.edu>
- ~25,000 STSs Mapped on a Genome-Wide Basis by YAC-Based STS-Content Mapping, Radiation Hybrid (RH) Mapping, and/or Genetic Mapping
- Strategy for Map Construction

Genetic and Radiation Hybrid Maps Provide Global Framework ('Top-Down Mapping')

YAC-Based STS-Content Map Provides Local Ordering of STSs ('Bottom-Up Mapping')

Cross-Reference the Top-Down and Bottom-Up Maps to Deduce an 'Integrated Map' of Each Chromosome

# The Whitehead Map: Data Generation

- ~25,000 STSs Mapped

Random Sequences	21%
Genetic Markers	28%
Expressed Sequences (Genes/ESTs)	51%

- Integrated Approach for Physical Mapping of STSs

## 1. YAC-Based STS-Content Mapping

Detect Linkage Over Distance of ~1 Mb

~11,000 STSs

CEPH Mega-YACs

## 2. Radiation Hybrid Mapping of STSs

Detect Linkage Over Distance of ~10 Mb

~15,000 STSs

GeneBridge 4 Whole Genome Radiation Hybrid Panel

## 3. Genethon Genetic Maps

Detect Linkage Over Distance of ~30 Mb

~5,300 STSs

Note: Partial Overlap Among STSs (~25,000 STSs Total)

- PCR Analysis: *Genomatron*

Massively-Parallel, Factory-Style Automation System

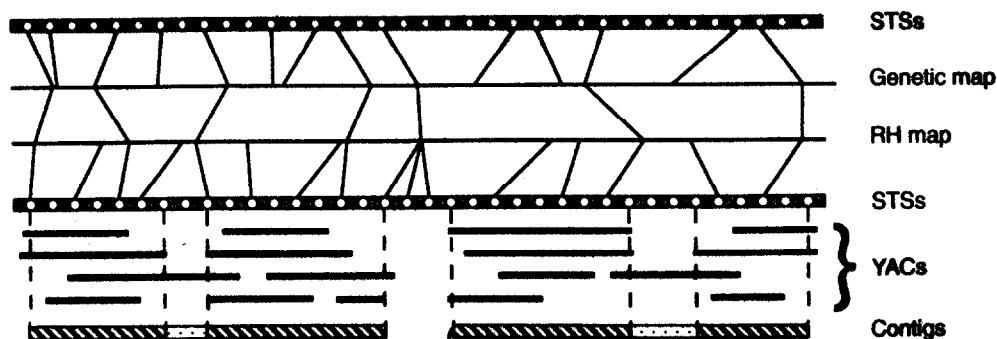
1536 Position Arrays

Highly Integrated for PCR Setup and Analysis

~150,000 PCR Assays Per Run

# The Whitehead Map: Data Analysis

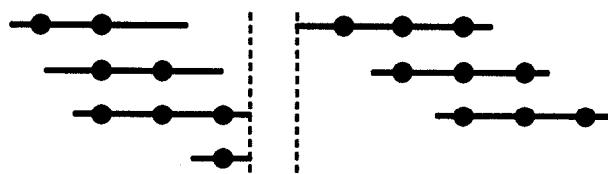
- Integration of Top-Down and Bottom-Up Maps



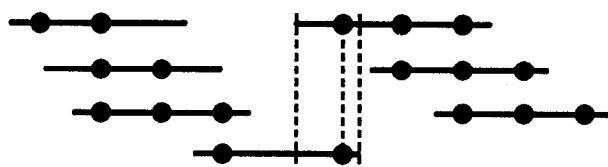
1. YAC-Based STS-Content Mapping
2. Assemble “Double-Linked” YAC Contigs
3. Evaluate Proximity to “Nearby” YAC Contigs
4. Allow “Single YAC Links” to “Appropriate” Nearby Contigs
5. Close Gaps by Use of CEPH YAC Data

- Closure Issues

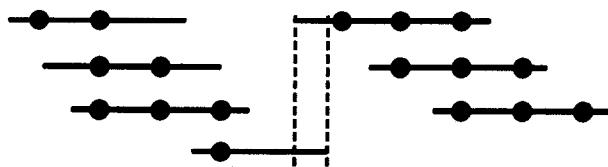
## A. Gap



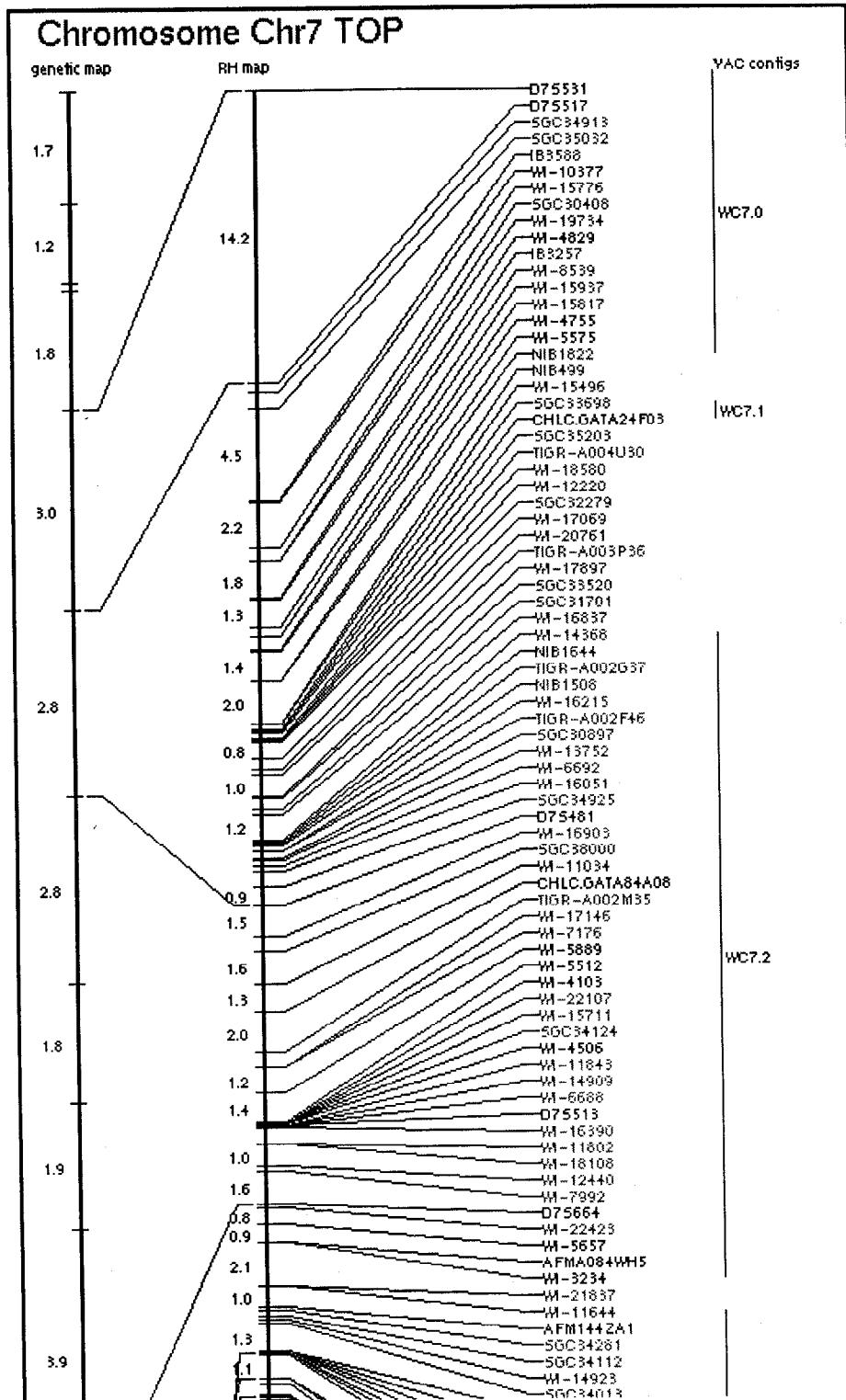
## B. Overlap



## C. Unknown Overlap



# The Whitehead Map: An Example



## **The Whitehead Map: Relevant Highlights**

- Average STS Resolution: ~120 kb
- Estimated Coverage of Human Genome: 94%
- Total Amount of PCR Performed: >25,000,000 PCR Assays
- Only Partially a Clone-Based Physical Map
- Limited Information about Inter-STS Distances
- Excellent Availability of Data and Maps

<http://www-genome.wi.mit.edu>

- Ability of Investigators to Map New STSs:

CEPH Mega-YACs  
Radiation Hybrid Cell Line DNA  
Appropriate Informatics Tools

# **Physical Mapping of the Mouse and Rat Genomes**

- **Mouse**

Among the ‘Sanctioned’ Organisms in the Genome Project

YAC and BAC Libraries Available

Major Mapping Effort at Whitehead Institute

Completion of Genetic Map Containing ~7,400 Markers

~8,000 STSs Screened Against a YAC Library (To Date)

Plan to Map ~10,000 STSs Against YACs

Similar Data Analysis as for Constructing the Human Map  
(Except Lack RH Mapping Data)

- **Rat**

Not Among the ‘Sanctioned’ Organisms in the Genome Project

Modest Effort at Several Locations (Including Whitehead Institute)

Major Emphasis to Date on Constructing Genetic Map  
(~2,000 Markers to Date)

YAC and BAC Libraries Available

Unclear if Clone-Based vs. RH Mapping for Rat Physical Map

## **WWW Access to Human Physical Maps and Data**

- Individual Web Sites Associated with Genome-Wide and Chromosome-Specific Projects (Listed Throughout Handout)

- Public Databases:

GenBank/Entrez: <http://www.ncbi.nlm.nih.gov>

GDB: <http://gdbwww.gdb.org>

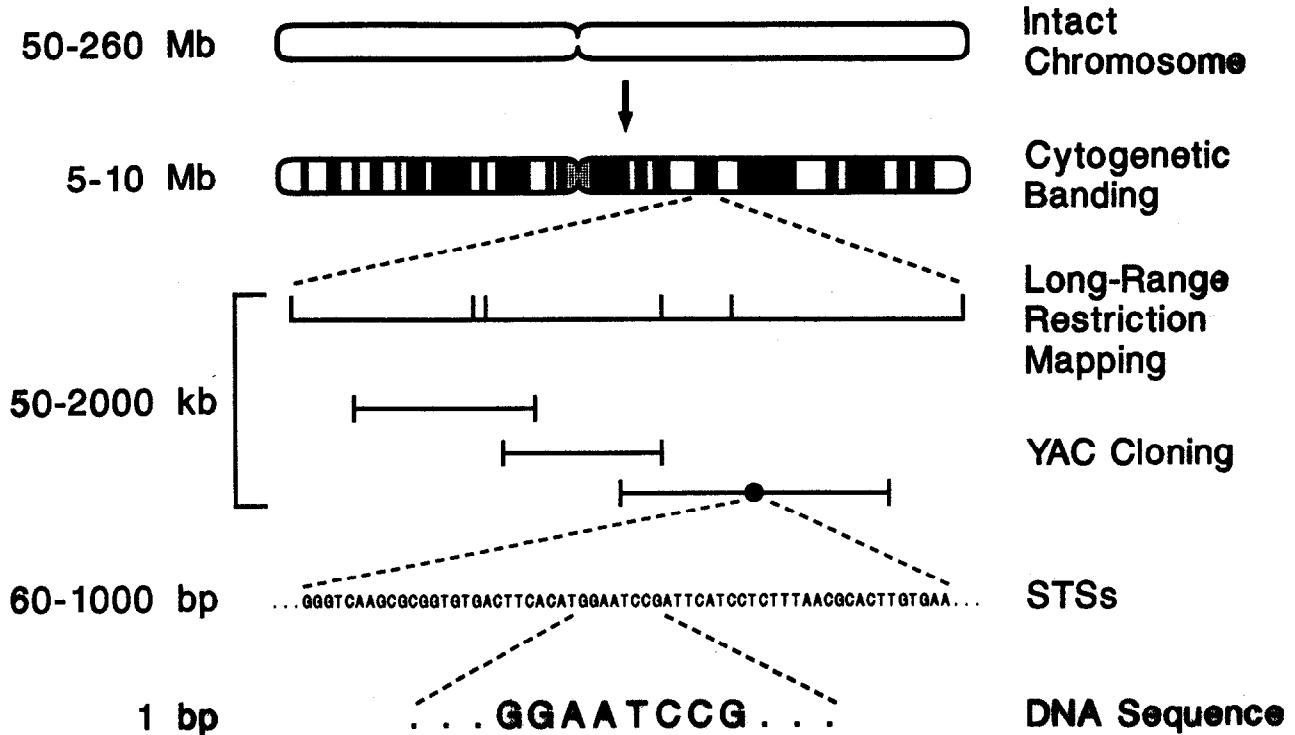
- ‘Guides’ to Numerous Other Relevant WWW Sites:

NHGRI Web Site: <http://www.nhgri.nih.gov/Data/>

GDB Web Site: [http://gdbwww.gdb.org/gdb/  
hgpResources.html](http://gdbwww.gdb.org/gdb/hgpResources.html)

## Future Prospects: 'Onward to Sequencing'

### Resolution Range:



- Olson (1995)

## A Time to Sequence

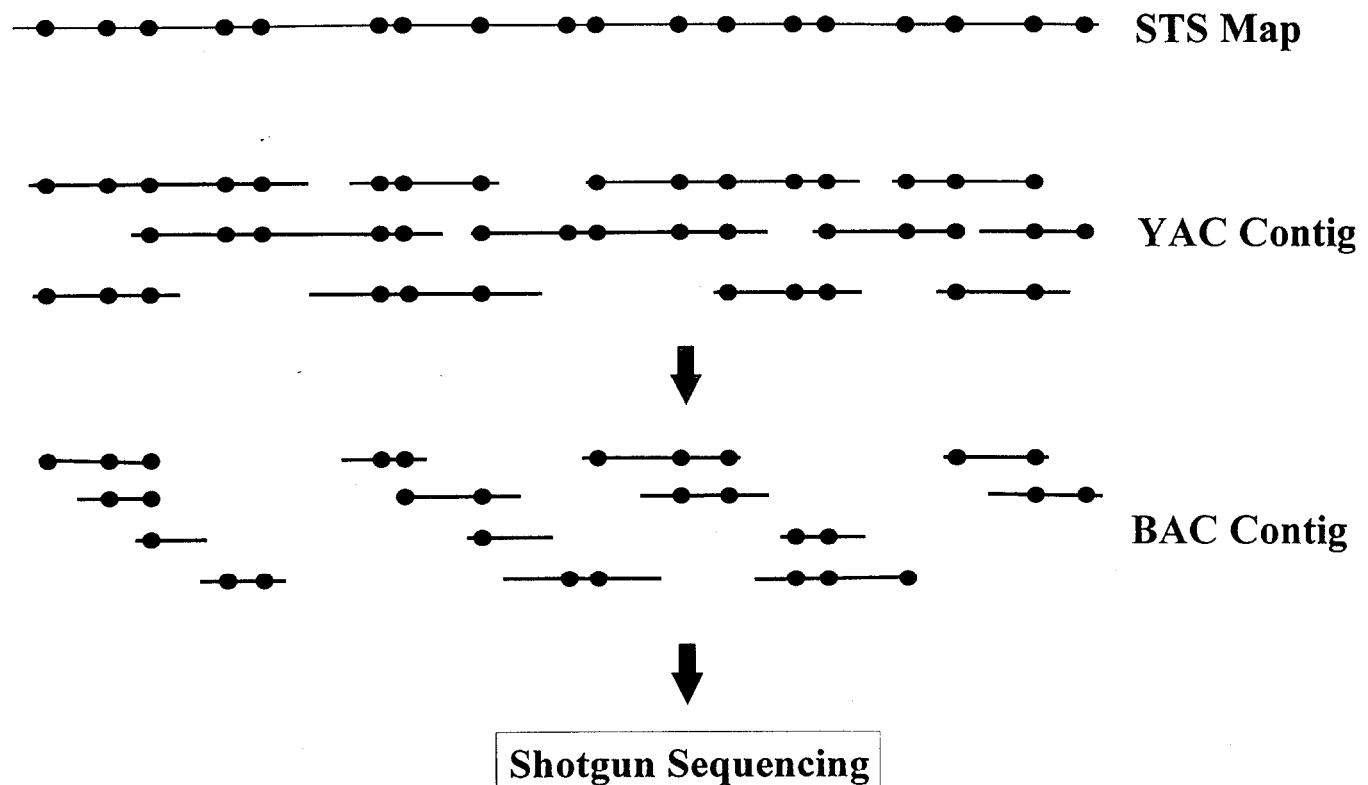
Maynard V. Olson

- Boguski et al. (1996)

### EDITORIAL

## The End of the Beginning: The Race to Begin Human Genome Sequencing

# **Construction of ‘Sequence-Ready Maps’**



## BIBLIOGRAPHY

- Albertsen, H.M. et al (1990). Construction and characterization of a yeast artificial chromosome library containing seven haploid genome equivalents. *Proc. Natl. Acad. Sci. USA* 87, 4256-4260.
- Anand, R. et al (1989). Construction of yeast artificial chromosome libraries with large inserts using fractionation by pulsed-field gel electrophoresis. *Nucl. Acids Res.* 17, 3425-3433.
- Anand, R. et al (1990). A 3.5 genome equivalent multi access YAC library: construction, characterisation, screening and storage. *Nucl. Acids Res.* 18, 1951-1956.
- Arratia, R. et al (1991). Genomic mapping by anchoring random clones: a mathematical analysis. *Genomics* 11, 806-827.
- Ashworth, L.K. et al (1995). An integrated metric physical map of human chromosome 19. *Nature Genetics* 11, 422-427.
- Ballabio, A. (1993). The rise and fall of positional cloning. *Nature Genetics* 3, 277-279.
- Barillot, E. et al (1991). Theoretical analysis of a physical mapping strategy using random single-copy landmarks. *Proc. Natl. Acad. Sci. USA* 88, 3917-3921.
- Bell, C.J. et al (1995). Integration of physical, breakpoint and genetic maps of chromosome 22. Localization of 587 yeast artificial chromosomes with 238 mapped markers. *Hum. Mol. Genet.* 4, 59-69.
- Bellanne-Chantelot, C. et al (1992). Mapping the whole human genome by fingerprinting yeast artificial chromosomes. *Cell* 70, 1059-1068.
- Bentley, D.R. et al (1992). The development and application of automated gridding for efficient screening of yeast and bacterial ordered libraries. *Genomics* 12, 534-541.
- Boguski, M. et al (1996). The end of the beginning: the race to begin human genome sequencing. *Genome Research* 6, 771-772.

- Bouffard, G.G. et al (1997). A physical map of human chromosome 7: an integrated YAC contig map with average STS spacing of 79 kb. *Genome Research* 7, 673-692.
- Brownstein, B.H. et al (1989). Isolation of single-copy human genes from a library of yeast artificial chromosome clones. *Science* 244, 1348-1351.
- Burke, D.T. et al (1987). Cloning of large segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 236, 806-812.
- Burke, D.T. and Olson, M.V. (1991). Preparation of clone libraries in yeast artificial-chromosome vectors. In *Methods in Enzymology* (Vol. 194): Guide to Yeast Genetics and Molecular Biology. C. Guthrie and G.R. Fink, eds. (San Diego: Academic Press, Inc.), pp. 251-270.
- Burke, D.T. et al (1991). A mouse genomic library of yeast artificial chromosome clones. *Mammalian Genome* 1, 65.
- Burke, D.T. (1991). The role of yeast artificial chromosome clones in generating genome maps. *Curr. Opin. Genet. Develop.* 1, 69-74.
- Cai, L. et al (1997). Construction and characterization of a 10-genome equivalent yeast artificial chromosome library for the laboratory rat, *Rattus norvegicus*. *Genomics* 39, 385-392.
- Chartier, F.L. et al (1992). Construction of a mouse yeast artificial chromosome library in a recombination-deficient strain of yeast. *Nature Genet.* 1, 132-136.
- Chumakov, I. et al (1992). Continuum of overlapping clones spanning the entire human chromosome 21q. *Nature* 359, 380-387.
- Chumakov, I.M. et al (1995). A YAC contig map of the human genome. *Nature* 377, S175-S298.
- Cohen, D. et al (1993). A first-generation physical map of the human genome. *Nature* 366, 698-701.
- Collins, F.S. (1992). Positional cloning: let's not call it reverse anymore. *Nature Genetics* 1, 3-6.
- Collins, F. and Galas, D. (1993). A new five-year plan for the U.S. human genome project. *Science* 262, 43-46.

Collins, F.S. (1995). Positional cloning moves from the perditional to traditional. *Nature Genetics* 9, 347-350.

Collins, J.E. et al (1995). A high-density YAC contig map of human chromosome 22. *Nature* 377, S367-S379.

Coulson, A. et al (1986). Toward a physical map of the genome of the nematode *Caenorhabditis elegans*. *Proc. Natl. Acad. Sci. USA* 83, 7821-7825.

Dausset, J. et al (1992). The CEPH YAC library. *Behring Inst. Mitt.* 91, 13-20.

Doggett, N.A. et al (1995). An integrated physical map of human chromosome 16. *Nature* 377, S335-S366.

Evans, G.A. et al (1992). Use of cosmids and arrayed clone libraries for genome analysis. *Methods in Enzymology* 216, 530-548.

Foote, S. et al (1992). The human Y chromosome: overlapping DNA clones spanning the euchromatic region. *Science* 258, 60-66.

Gemmill, R.M. et al (1995). A second-generation YAC contig map of human chromosome 3. *Nature* 377, S299-S320.

Goold, R.D. et al (1993). The development of sequence-tagged sites for human chromosome 4. *Hum. Mol. Genet.* 2, 1271-1288.

Green, E.D. and Olson, M.V. (1990). Systematic screening of yeast artificial-chromosome libraries by use of the polymerase chain reaction. *Proc. Natl. Acad. Sci. USA* 87, 1213-1217.

Green, E.D. and Olson, M.V. (1990). Chromosomal region of the cystic fibrosis gene in yeast artificial chromosomes: a model for human genome mapping. *Science* 250, 94-98.

Green, E.D. et al (1991). Detection and characterization of chimeric yeast artificial-chromosome clones. *Genomics* 11, 658-669.

Green, E.D. and Green, P. (1991). Sequence-tagged site (STS) content mapping of human chromosomes: theoretical considerations and early experiences. *PCR Methods Appl.* 1, 77-90.

Green, E.D. et al (1994). Integration of physical, genetic and cytogenetic maps of human chromosome 7: isolation and analysis of yeast artificial chromosome clones for 117 mapped genetic markers. *Hum. Mol. Genet.* 3, 489-501.

Green, E.D. et al (1995). A human chromosome 7 yeast artificial chromosome (YAC) resource: construction, characterization, and screening. *Genomics* 25, 170-183.

Haldi, M.L et al (1996). A comprehensive large-insert yeast artificial chromosome library for physical mapping of the mouse genome. *Mammalian Genome* 7, 767-769.

Haldi, M.L et al (1997). Construction of a large-insert yeast artificial chromosome library of the rat genome. *Mammalian Genome* 8, 284.

Haldi, M.L et al (1997). Construction of a large-insert yeast artificial chromosome library of the rat genome. *Mammalian Genome* 8, 460.

Hieter, P. et al (1990). Yeast artificial chromosomes: promises kept and pending. In *Genome Analysis Volume 1: Genetic and Physical Mapping*. K.E. Davies and S.M. Tilghman, eds. (Cold Spring Harbor: Cold Spring Harbor Laboratory Press), pp. 83-120.

Hudson, T.J. et al (1995). An STS-based map of the human genome. *Science* 270, 1945-1954.

Ioannou, P.A. et al (1994). A new bacteriophage P1-derived vector for the propagation of large human DNA fragments. *Nature Genet.* 6, 84-89.

Ivens, A.C. et al (1993). Cosmid clones and generation of detailed DNA maps of large chromosomal regions. In *Methods in Molecular Genetics (Gene and Chromosome Analysis, Part A)*. K.W. Adolph, ed. (San Diego: Academic Press), pp. 173-191.

Kohara, Y. et al (1987). The physical map of the whole *E. coli* chromosome: application of a new strategy for rapid analysis and sorting of a large genomic library. *Cell* 50, 495-508.

Krauter, K. et al (1995). A second-generation YAC contig map of human chromosome 12. *Nature* 377, S321-S334.

Kusumi, K. et al (1993). Construction of a large-insert yeast artificial chromosome library of the mouse genome. *Mammalian Genome* 4, 391-392.

Larin, Z. et al (1991). Yeast artificial chromosome libraries containing large inserts from mouse and human DNA. *Proc. Natl. Acad. Sci. USA* 88, 4123-4127.

Larin, Z. et al (1993). Construction and characterization of yeast artificial chromosome libraries from the mouse genome. *Methods in Enzymology* 225, 623-637.

Nagaraja et al (1997). X chromosome map at 75-kb STS resolution, revealing extremes of recombination and GC content. *Genome Research* 7, 210-222.

Nelson, D.L. et al (1989). *Alu* polymerase chain reaction: a method for rapid isolation of human-specific sequences from complex DNA sources. *Proc. Natl. Acad. Sci. USA* 86, 6686-6690.

Nelson, D.L. (1991). Interspersed repetitive sequence polymerase chain reaction (IRS PCR) for generation of human DNA fragments from complex sources. *Methods: A Companion to Methods in Enzymology* 2, 60-74.

Nelson, D.L. et al (1991). *Alu*-primed polymerase chain reaction for regional assignment of 110 yeast artificial chromosome clones from the human X chromosome: identification of clones associated with a disease locus. *Proc. Natl. Acad. Sci. USA* 88, 6157-6161.

Nizetic, D. et al (1991). Construction, arraying, and high-density screening of large insert libraries of human chromosomes X and 21: their potential use as reference libraries. *Proc. Natl. Acad. Sci. USA* 88, 3233-3237.

Olsen, A.S. et al (1993). Automated production of high density cosmid and YAC colony filters using a robotic workstation. *BioTechniques* 14, 116-123.

Olson, M.V. et al (1986). Random-clone strategy for genomic restriction mapping in yeast. *Proc. Natl. Acad. Sci. USA* 83, 7826-7830.

Olson, M. et al (1989). A common language for physical mapping of the human genome. *Science* 245, 1434-1440.

Olson, M.V. and Green, P. (1993). Criterion for the completeness of large-scale physical maps of DNA. *Cold Spring Harb. Symp. Quant. Biol. LVIII*, 349-355.

Olson, M.V. (1995). A time to sequence. *Science* 270, 394-396.

Palazzolo, M.J. et al (1991). Optimized strategies for sequence-tagged-site selection in genome mapping. *Proc. Natl. Acad. Sci. USA* 88, 8034-8038.

Pierce, J.C. and Sternberg, N. (1992). The bacteriophage P1 cloning system. In Techniques for the analysis of complex genomes. R. Anand, ed. (London: Academic Press Limited), pp. 39-58.

Pierce, J.C. and Sternberg, N.L. (1992). Using bacteriophage P1 system to clone high molecular weight genomic DNA. *Methods in Enzymology* 216, 549-574.

Qin, S. et al (1996). A high-resolution physical map of human chromosome 11. *Proc. Natl. Acad. Sci.* 93, 3149-3154.

Quackenbush, J. et al (1995). An STS content map of human chromosome 11: localization of 910 YAC clones and 109 islands. *Genomics* 29, 512-525.

Ramsay, M. and Wicking, C. (1991). Yeast artificial-chromosome (YAC) cloning systems. In *Protocols in Human Molecular Genetics (Methods in Molecular Biology, Vol. 9)*. C.G. Mathew, ed. (Clifton, NJ: Humana Press, Inc.), pp. 197-221.

Riles, L. et al (1993). Physical maps of the six smallest chromosomes of *Saccharomyces cerevisiae* at a resolution of 2.6 kilobase pairs. *Genetics* 134, 81-150.

Ross, M.T. et al (1992). High-density gridded YAC filters: their potential as genome mapping tools. In Techniques for the analysis of complex genomes. R. Anand, ed. (London: Academic Press Limited), pp. 137-153.

Rossi, J.M. et al (1992). Genomic analysis using a yeast artificial chromosome library with mouse DNA inserts. *Proc. Natl. Acad. Sci. USA* 89, 2456-2460.

Sambrook, J. et al (1989). *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor: Cold Spring Harbor Laboratory Press).

Schlessinger, D. (1990). Yeast artificial chromosomes: tools for mapping and analysis of complex genomes. *Trends Genet.* 6, 248-258.

Schlessinger, D. and Kere, J. (1992). YAC-based mapping of genome structure, function, and evolution. In *Genome Analysis Volume 4: Strategies for Physical*

Mapping. K.E. Davies and S.M. Tilghman, eds. (Cold Spring Harbor: Cold Spring Harbor Laboratory Press), pp. 131-159.

Shepherd, N.S. et al (1994). Preparation and screening of an arrayed human genomic library generated with the P1 cloning system. Proc. Natl. Acad. Sci. USA 91, 2629-2633.

Shizuya, H. et al (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. Proc. Natl. Acad. Sci. USA 89, 8794-8797.

Smith, M.W. et al (1993). A sequence-tagged site map of human chromosome 11. Genomics 17, 699-725.

Stallings, R.L. et al (1990). Physical mapping of human chromosomes by repetitive sequence fingerprinting. Proc. Natl. Acad. Sci. USA 87, 6218-6222.

Sternberg, N. (1990). Bacteriophage P1 cloning system for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs. Proc. Natl. Acad. Sci. USA 87, 103-107.

Sternberg, N. et al (1990). Generation of a 50,000-member human DNA library with an average DNA insert size of 75-100 kbp in a bacteriophage P1 cloning vector. The New Biologist 2, 151-162.

Vollrath, D. et al (1992). The human Y chromosome: a 43-interval map based on naturally occurring deletions. Science 258, 52-59.

Wahl, G.M. et al (1987). Cosmid vectors for rapid genomic walking, restriction mapping, and gene transfer. Proc. Natl. Acad. Sci. USA 84, 2160-2164.